

Towards a Hybrid Space Combining Spatial Augmented Reality and Virtual Reality

Joan Sol Roo ^{*}
Inria, France

Martin Hachet [†]
Inria, France

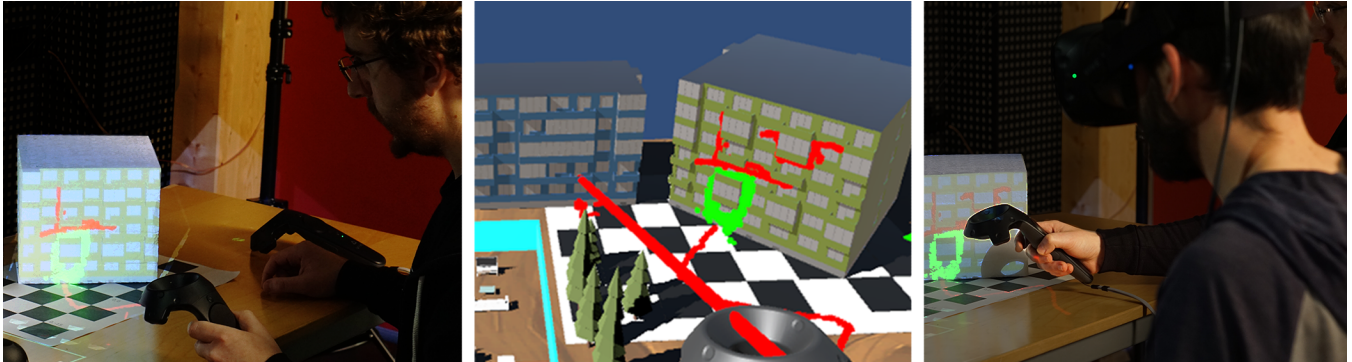


Figure 1: A coherent interaction with an augmented space, combining Spatial Augmented Reality and an immersive head mounted display (HMD). Left: a user interacts with an augmented mock-up. Center: wearing the HMD's enables to see virtual elements. Right: two users can interact and collaborate each using a different modality.

ABSTRACT

Spatial Augmented Reality (SAR) allows a user, or a group of users, to benefit from digital augmentations embedded directly into the physical world. This enables co-located information and unobstructed interaction. On the other hand, SAR suffers from limitations that are inherently linked to its physical dependency, which is not the case for see-through or immersive displays. In this work, we explore how to facilitate the transition from SAR to VR, and vice versa, integrating both into a unified experience. We developed a set of interaction techniques and obtained first feedback from informal interviews.

Keywords: Mixed Reality, Augmented Reality, HMD, 3D Interaction

Index Terms: H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities—; H.5.2 [User Interfaces]: Input devices and strategies—

1 INTRODUCTION

Spatial Augmented Reality (SAR) mixes physical and digital information directly onto the real world by the mean of screens or projectors [19]. SAR leverages our already available skills to interact with the physical world. One or more users can interact with the same coherent illusion, without the obstruction of instrumentation (Figure 1, left).

Despite its huge potential, SAR has also limitations. The used technology may penalize the immersion of the user in the hybrid world (e.g., the light path of the projectors may generate occlusions, and the resulting image will depend on the projector brightness and

the surface material). Moreover, the need of a physical support to display digital information is a strong constraint of SAR. Indeed, it is not possible to project mid air information with the current technology, unless using anamorphic illusions [3]. Even when this limitation is partially overcome by see-through augmented reality (ST-AR) devices such as Microsoft HoloLens, they are still constrained by the physical world, with the additional requirement to instrument the user. Alternatively, immersive head mounted displays for Virtual Reality (HMD-VR) allow us to override completely the perceived reality; with HMD-VR it is possible to freely interact with the perceived scene, moving and scaling objects or changing the point of view independently from the physical space. There is then a clear trade-off between instrumenting the user to increase the quality of the illusion, and keeping the interaction unencumbered.

An example that benefits from mixed reality are mock-ups, physical replicas used to preview artifacts or buildings. Thanks to SAR, these physical objects can be augmented with dynamic content, albeit with limited resolution and without mid air information (Figure 1). This is particularly interesting for prototyping [15], or for a group of persons to perceive how the final project would look like, and to discuss the possible improvements. The problem with this approach is that the scene cannot be scaled or navigated. On the other hand, HMD-VR allow users to perform these interactive tasks [13], at the cost of losing the physical properties. Combining multiple mixed reality modalities would complement their individual weaknesses.

Our motivation for this work is to explore how SAR and HMD-VR technologies can complement each other. We aim at bridging the gap between these two approaches that have tended to evolve in separate tracks. More precisely, we explore how to favor transitions between SAR and VR by proposing unified interaction and several awareness helpers.

2 RELATED WORK

The notion that the physical and digital worlds can be combined in varying degrees is not new, first formalized in 1994 by Milgram et

^{*}e-mail: joan-sol.roo@inria.fr

[†]martin.hachet@inria.fr

al. with the virtuality-reality continuum [16]. This section briefly describes previous work that focused on enabling the combination and transition between mixed realities.

See-through displays and SAR have been combined in the past, notably in order to complement the HMD's high resolution with the projectors' scalable FOV (field of view) [10, 2]. In the context of multi-display environments, the combination of screens and projection has been studied, both with [5] and without see-through technologies [20, 7]. Other hybrid mixed reality systems have also been explored. For example, Magicbook [4] combines a physical non-augmented book with video see-through, enabling the transition from non augmented interaction, through ST-AR to immersive VR. Dedual et al. [6] proposed a system that combines interactive tabletops with video see-through AR. Smarter Objects [8] and ex-Touch [12] use video ST-AR to control embedded systems; even when the physical artifact was the focus of attention, no spatial augmentation was presented, except the electronic behaviour itself. Closer to our work is Metadesk [21] by Ullmer and Ishii, which combines tabletop interaction, tangible tokens and a see-through AR window.

The aforementioned projects are based on complementary technologies, combined in a unified frame of reference, where the scene is observed from a unique physical location. Compared to these works, we are interested in interaction metaphors where the virtual component is used to explore different scales and points of view, without being restricted to the user's location. On this direction, Ibayashi et al. created Dollhouse VR [9], where users can interact with a virtual space, either with a top view (displayed on a tabletop), or from inside (using a HMD-VR). Our work extends Dollhouse VR by supporting a non-flat physical support (using SAR instead of a tabletop), while also providing a see-through modality between physical and immersed.

Our work complements the state of the art by supporting transitions between SAR and HMD-VR, using a physical artifact as starting point while also enabling free navigation of the virtual scene.

3 SAR/VR HYBRID SPACE

We propose a hybrid SAR/VR hybrid space that enables both 1) the interaction of one or various users with an augmented object, and 2) immersive experiences. A unified view provides traceability of what happens in VR while in SAR, and vice versa. This can be used for symmetric and asymmetric collaboration between users, or to increase the perceived information for one user. This section describes the interaction techniques for one user, yet multiple users can use the system at the same time, their activities visible to one another thanks to the considerations presented on Section 3.7.

3.1 Scenario

As an example, let us consider two users interacting with the hybrid mock-up. They can explore it by simply physically moving around, and interact with their hands (touching or pointing with their finger). They can also use an augmented tool to do annotations directly onto the mock-up. Now, one user wants to push the frontiers of physicality. By wearing a VR helmet, she can observe additional information (e.g. the virtual landscape surrounding the mock-up). She can also decide to navigate inside the scene to observe the mock-up from an egocentric point of view. Navigating virtually around the mock-up with an exocentric point of view, the user can change her perception of the scene by observing it from a location that would not be accessible with a standard AR approach. A scanned version of both users can be displayed to reinforce the link with the physical world. The navigation can involve both a change in position and scale. At all times, any of the users can annotate the mock-up, and the annotation will be visible to the other user. For awareness reasons, if one of the users is immersed in the mock-up, the position is indicated on the surface.

While in VR, it is also possible to use traditional virtual interaction, complementing classical limitations of the physical world. Virtual objects can be translated, rotated and scaled, and the internal components exposed. These changes can be reflected on the augmented mock-up to different extents, depending on the available physical geometry.

3.2 Scene Elements

In order to understand the similarities and differences on how a user interacts with the mock-up, it is necessary to mention that an element on the scene can be either: physical, virtual, or augmented (both physical and virtual).

- **Physical** elements are already available in reality and do not need to be displayed in SAR
- **Virtual** elements can be seen in VR, but they are not readily available in SAR. Nevertheless, they can be displayed as simplified augmented surfaces (for example, generating a 2d projection of a 3d object)
- **Augmented** elements have dual physical-virtual characteristics, and they can be interacted with differently in SAR and VR

3.3 Unified Interaction

In order to support a unified interaction with both SAR and VR, we decided to implement all the interaction techniques using a 6DoF wand controller.

3.4 Pointing and Annotation

Most interactions with the scene require to be able to point. This is solved via ray-casting, using the controller's forward direction, with an inclination of 55 downwards to ease the interaction with nearby objects. While in VR, the cursor is represented with a ray; the SAR modality only displays the target cursor, since it is not possible to show the ray in mid air.

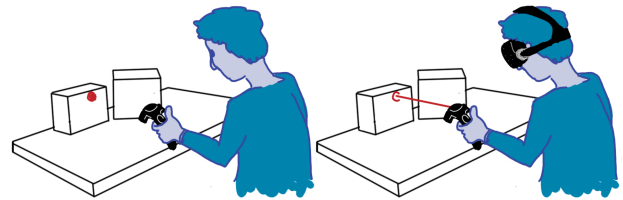


Figure 2: Pointing on the mock-up using a wand controller. Left: pointing in SAR shows a target directly onto the surface. Right: pointing in VR shows a ray.

In order to perform free hand annotations on a surface, the user can simply point at the target location and press the controller's trigger to start drawing (Figure 5).

3.5 Manipulation

Physical objects can be manipulated with the bare hands, and augmented versions can be manipulated when tracked [14]. Virtual objects can also be manipulated using the controller's grab button. This action is best suited for VR, but it is possible to still perform it in SAR. If the virtual object has a physical counterpart (i.e., they form an augmented object), they will leave behind a ghost when grabbed, and the virtual model will return to the default physical position when dropped, in order to reinforce the illusion. Purely virtual objects can be released in a new position freely.

Bimanual operation is also possible, using one controller to change the position and orientation of the virtual object, and a second one to create annotations. While in SAR, both augmented and virtual objects will reflect the annotations, yet augmented objects will not change position.

3.6 Navigation

The physical mock-up can be explored by naturally navigating the physical space, or by virtually changing the point of view (Figure 3).

- **Jump in:** A novel point of view can be selected by pointing at the desired surface, either on SAR or VR, and then pressing the controller's trackpad. This will teleport the user to the selected position on the surface, while also changing the user's scale to a fixed factor (depending on the mock-up characteristics). While on the surface, the trackpad can be used to adjust the orientation. The user can move by selecting a new jump position. Note that teleportation is the standard navigation technique for HTC Vive, selected to prevent cyber-sickness.
- **Orbit:** While in VR and outside the mock-up, it is possible to orbit around the table, using the trackpad to indicate the angle.
- **Reset navigation:** At any time, the user can point the controller up and press the trackpad to return to the natural POV. Resetting the navigation before removing the HMD eases the transition back to reality.

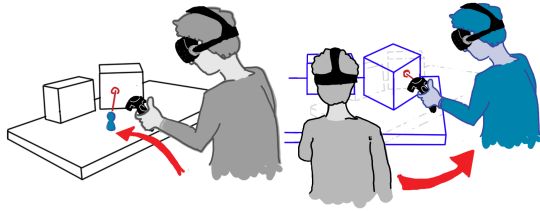


Figure 3: Detaching the point of view from the physical location. The user in gray represents the physical location, while the blue user indicates the perceived location. Left: traveling inside the mock-up. Right: orbiting around the table.

3.7 Awareness

In order to prevent the users to feel lost once they change their point of view, the physical surroundings of the mock-up –including the users– are reconstructed in 3D (Figure 4).



Figure 4: The user orbited left, so he can see a 3D reconstruction of himself and the chairs when looking to his right.

To ease the understanding of where the user is located when they "jump in" the mock-up, an arrow is displayed on the SAR version of the system, indicating their position and orientation (Figure 5). This, in combination of the visible cursor enables non-immersed users to communicate with the immersed ones.



Figure 5: Left: the user is inside the virtual room looking at a building. Right: the user's position and orientation are indicated with a yellow arrow in SAR. The annotation over the augmented object is visible in SAR, while the annotation over the virtual wall is presented as a top-down flat projection.

Navigation creates a disconnection between the physical environment and what the user sees, which can lead to accidents. To prevent this, real objects are shown as ghosts when getting close to them (Figure 6), similar to the HTC Vive approach safety cage.

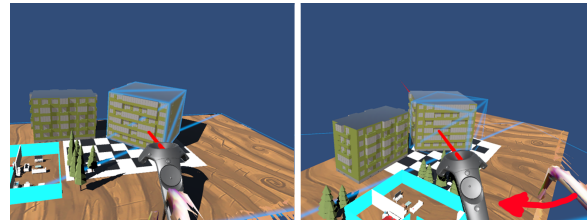


Figure 6: If the controller approaches physical objects (in this case, the table and the right building), their bounding boxes are displayed. Left: since the user did not change the point of view, then the bounding box matches the perceived virtual element. Right: the user performed an orbit operation to the left, yet the bounding box is still displayed in the same physical location.

4 IMPLEMENTATION

The main application was developed using Unity3D 5.4 and DirectX 11. The software runs on an Alienware i7-3820 with 8GiB of RAM and a GeForce GTX 660. The setup is comprised of: a Microsoft Kinect v1, a LG 1500G projector and a HTC Vive with its controllers. The overall system runs at 60 frames per second, without particular optimization considerations, except by the processing of the Kinect data using the GPU.

The calibration of the optical devices was performed offline, setting the center of the table as origin. The projector calibration uses OpenCV's camera calibration functionality (2D to 3D points), while the depth camera was calibrated using Kabsch algorithm (3D to 3D points, rotation and translation). The alignment of HTC Vive with our system's frame of reference is performed online. The transformation between coordinate systems is computed as the inverse of the controller's origin at a known table position (i.e., resting position over the table's origin).

4.1 Limitations

The current implementation could be greatly improved using the advances of avatar generation in the area of tele-presence [18, 1]. Object tracking is not currently supported, yet it can be implemented using off the shelf solutions such as Optitrack (taking into account infrared interference from the current sensors), or marker-less tracking [17]. Free hand interaction while in VR could also be supported, using consumer grade sensors such as Leap Motion. The main objective of this project was not to exhaustively explore these alternatives, but instead to take the first steps towards the use of immersive displays to support spatially augmented scenes.

5 INTERVIEWS

We conducted informal interviews with 3 colleagues from related fields to know more about the acceptability and viability of the hybrid space. We presented the system features incrementally, and discussed their perspective on the positive aspects as well as the limitations of the system. The collected comments are briefly summarized in this section.

- **Participant 1 (tangible interfaces and SAR):** he was interested on the flexibility of the system in contrast with pure tangible interfaces, specially the change in perspective, while not completely leaving the physical world. He suggested to include support for direct interaction.
- **Participant 2 (collaborative VR):** he was particularly pleased with the feeling of presence once inside the mock-up. He suggested to support mid air annotations while in VR, and also recommended to track the user's hands, using for example skeleton tracking.
- **Participant 3 (Spatial navigation and maps):** she found that the mock-up was a good example of semantic zoom, with the physical props working as low resolution landmarks. Nevertheless, she mentioned that the translation action was too abrupt, and that it would benefit from a smoother transitioning.

Even when these interviews did not validate the system, they showed the system can be promising for different fields, yet further work would be required.

6 CONCLUSIONS

In this work we presented a hybrid space comprised of Spatial augmented reality augmentation and immersive displays. Interaction techniques were presented to annotate and interact with the scene, along with the navigation of the virtual counterpart. The discussions with practitioners in related fields show that merging VR and SAR is promising, yet there is space for improvement by adopting additional interaction techniques, such as direct interaction or skeleton tracking.

In this work we took the example of mock-ups in the field of architecture. Of course, this approach extends to many scenarios, including object design, physical visualization [11], or education.

The perception of this hybrid space as a unique world is a key requirement for its success. Future work will involve perceptual user studies, which will help us to improve this hybrid space and to make the link between the two worlds stronger; then, it would be of interest to study the interaction between multiple users in collaborative scenarios. These studies will build towards the creation of novel hybrid applications.

7 ACKNOWLEDGEMENTS

This work was supported by the ANR-ISAR project, ANR-14-CE24-0013.

REFERENCES

- [1] S. Beck, A. Kunert, A. Kulik, and B. Froehlich. Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):616–625, 2013.
- [2] H. Benko, E. Ofek, F. Zheng, and A. D. Wilson. Fovear: Combining an optically see-through near-eye display with projector-based spatial augmented reality. In *UIST '15*, pages 129–135. ACM, 2015.
- [3] H. Benko, A. D. Wilson, and F. Zannier. Dyadic projected spatial augmented reality. In *UIST '14*, pages 645–655. ACM, 2014.
- [4] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook-moving seamlessly between reality and virtuality. *IEEE CGA*, 21(3):6–8, 2001.
- [5] A. Butz, T. Hollerer, S. Feiner, B. MacIntyre, and C. Beshers. Enveloping users and computers in a collaborative 3d augmented reality. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 35–44. IEEE, 1999.
- [6] N. J. Dedual, O. Oda, and S. K. Feiner. Creating hybrid user interfaces with a 2d multi-touch tabletop and a 3d see-through head-worn display. In *ISMAR 2011*, pages 231–232. IEEE, 2011.
- [7] R. Gervais, J. S. Roo, and M. Hachet. Tangible viewports: Getting out of flatland in desktop environments. In *TEI '16*, pages 176–184. ACM, 2016.
- [8] V. Heun, S. Kasahara, and P. Maes. Smarter objects: using ar technology to program physical objects and their interactions. In *CHI '13 Extended Abstracts*, pages 961–966. ACM, 2013.
- [9] H. Ibayashi, Y. Sugiura, D. Sakamoto, N. Miyata, M. Tada, T. Okuma, T. Kurata, M. Mochimaru, and T. Igarashi. Dollhouse vr: a multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 Emerging Technologies*, page 8. ACM, 2015.
- [10] A. Ilie, K.-L. Low, G. Welch, A. Lastra, H. Fuchs, and B. Cairns. Combining head-mounted and projector-based displays for surgical training. *Presence: Teleoperators and Virtual Environments*, 13(2):128–145, 2004.
- [11] Y. Jansen, P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian, and K. Hornbæk. Opportunities and challenges for data physicalization. In *CHI '15*, pages 3227–3236. ACM, 2015.
- [12] S. Kasahara, R. Niiyama, V. Heun, and H. Ishii. extouch: Spatially-aware embodied manipulation of actuated objects mediated by augmented reality. In *TEI '13, TEI '13*, pages 223–228, New York, NY, USA, 2013. ACM.
- [13] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana. Virtual object manipulation on a table-top ar environment. In *Augmented Reality, 2000.(ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pages 111–119. Ieee, 2000.
- [14] J. Laviolle and M. Hachet. Papart: interactive 3d graphics and multi-touch augmented paper for artistic creation. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, pages 3–6. IEEE, 2012.
- [15] M. R. Marner, B. H. Thomas, and C. Sandor. Physical-virtual tools for spatial augmented reality user interfaces. In *ISMAR*, volume 9, pages 205–206, 2009.
- [16] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Photonics for industrial applications*, pages 282–292. International Society for Optics and Photonics, 1995.
- [17] R. A. Newcombe, D. Fox, and S. M. Seitz. Dynamicfusion: Reconstruction and tracking of non-rigid scenes in real-time. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 343–352, 2015.
- [18] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou, et al. Holoportation: Virtual 3d teleportation in real-time. In *UIST '16*, pages 741–754. ACM, 2016.
- [19] R. Raskar, G. Welch, K.-L. Low, and D. Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *Rendering Techniques 2001*, pages 89–102. Springer, 2001.
- [20] J. Rekimoto and M. Saitoh. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *CHI '99*, pages 378–385. ACM, 1999.
- [21] B. Ullmer and H. Ishii. The metadesk: models and prototypes for tangible user interfaces. In *UIST '97*, pages 223–232. ACM, 1997.